Review Paper ■

Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience

Kenneth D. Mandl, MD, MPH, J. Marc Overhage, MD, PhD, Michael M. Wagner, MD, PhD, William B. Lober, MS, MD, Paola Sebastiani, PhD, Farzad Mostashari, MD, MSPH, Julie A. Pavlin, MD, MPH, Per H. Gesteland, MD, Tracee Treadwell, DVM, MPH, Eileen Koski, MPhil, Lori Hutwagner, MS, David L. Buckeridge, MD, MSc, Raymond D. Aller, MD, Shaun Grannis, MD

Abstract Syndromic surveillance refers to methods relying on detection of individual and population health indicators that are discernible before confirmed diagnoses are made. In particular, prior to the laboratory confirmation of an infectious disease, ill persons may exhibit behavioral patterns, symptoms, signs, or laboratory findings that can be tracked through a variety of data sources. Syndromic surveillance systems are being developed locally, regionally, and nationally. The efforts have been largely directed at facilitating the early detection of a covert bioterrorist attack, but the technology may also be useful for general public health, clinical medicine, quality improvement, patient safety, and research. This report, authored by developers and methodologists involved in the design and deployment of the first wave of syndromic surveillance systems, is intended to serve as a guide for informaticians, public health managers, and practitioners who are currently planning deployment of such systems in their regions.

■ **J Am Med Inform Assoc.** 2004;11:141–150. DOI 10.1197/jamia.M1356.

Bioterrorism preparedness has been the subject of concentrated national effort[1] that has intensified since the events of fall 2001.[2] In response to these events, the biomedical, public health, defense, and intelligence communities are developing new approaches to real-time disease surveillance in an effort to augment existing public health surveillance systems. New information infrastructure and methods to support timely detection and monitoring,[3–7] including the discipline of syndromic surveillance, are evolving rapidly. The term *syndromic surveillance* refers to methods relying on detection of clinical case features that are discernable before confirmed diagnoses are made. In particular, prior to the laboratory confirmation of an infectious disease, ill persons may exhibit behavioral patterns, symptoms, signs, or laboratory findings that can be tracked through a variety of data sources. If the attack involved anthrax, for example, a syndromic surveillance system might detect a surge in influenza-like illness, thus, providing an early warning and a tool for monitoring an ongoing crisis.

Affiliations of the authors: Children's Hospital Informatics Program, Division of Emergency Medicine, Center for Biopreparedness at Children's Hospital Boston, Children's Hospital Boston, Harvard Medical School, Boston, MA (KDM); Indiana University School of Medicine, Regenstrief Institute, Indianapolis, IN (JMO, SG); The Real-time Outbreak and Disease Laboratory, Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA (MMW); Department of Medical Education and Biomedical Informatics, School of Medicine, University of Washington, Seattle, WA (WBL); Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA (PS); Division of Epidemiology, New York City Department of Public Health, New York, NY (FM); Walter Reed Army Institute of Research, Silver Spring, MD (JAP); University of Utah and Intermountain Health Care, Salt Lake City, UT (PHG); Bioterrorism Preparedness and Response Program, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA (TT, LH); Quest Diagnostics Incorporated, Teterboro, NJ (EK); Palo Alto Veterans Health Care System, Palo Alto, CA, and Stanford Medical Informatics, Stanford University, Stanford, CA (DLB); Acute Communicable Diseases Unit, Los Angeles County Public Health, Los Angeles, CA (RDA).

Work on the manuscript was supported in part by funding from the National Library of Medicine (grants R01LM07677-01, 2 T15 LM07117-06, GO8 LM06625-01, and T15 LM/DE07059; contract N01-LM-9-3536; and training grants 2 T15 LM07117-06, 01-T15/LM-7124), the Agency for Healthcare Research and Quality (contracts 290-00-0020 and 290-00-0009), the Defense Advanced Projects Research Agency (contract F30602-01-2-0550), the Centers for Disease Control and Prevention (cooperative agreement number U90/CCU318753-01), the Alfred P. Sloan Foundation (Grant 2002-12-1), and the Canadian Institutes of Health Research. The authors gratefully acknowledge the contributions of Drs. Daniel Pollock, John Loonsk, and Michael D. Jones from the Centers for Disease Control and Prevention and Michael K. Martin from the Connecticut Hospital Association. The authors would like to thank Dasha Cohen of the American Medical Informatics Association for facilitating the meeting of the authors.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States government or the agencies listed above.

Correspondence and reprints: Kenneth D. Mandl, MD, MPH, Division of Emergency Medicine, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115; e-mail: <kenneth_mandl@harvard.edu>.

Received for publication: 03/05/03; accepted for publication: 09/28/03.

Unlike traditional systems that generally utilize voluntary reports from providers to acquire data, contemporary syndromic surveillance relies on an approach in which data are continuously acquired through protocols or automated routines. The real-time nature of these syndromic systems makes them valuable for bioterrorism-related outbreak detection, monitoring, and investigation. These systems augment the capabilities of the alert frontline clinician who, athough an invaluable resource for outbreak detection, is generally better at recognizing individual cases rather than patterns of cases over time and across a region. Syndromic surveillance technology may be useful not only for bioterrorism event detection, but also for general public health, clinical medicine, quality improvement, patient safety, and research. This report, authored by developers and methodologists involved in the design and deployment of the first wave of syndromic surveillance systems, is intended to serve as a guide for informaticians, public health managers, and practitioners who may be planning deployment of such systems in their regions.

Defining Leadership and Coalition

Participants who are necessary for establishing syndromic surveillance in a region include the originators of surveillance data (data providers) and a public health authority to receive and react to the data. In many cases, sufficient regional coverage may be achieved with data from a few large data providers. The coalition may also include "trusted brokers" (nonpartisan entities that receive and store data on behalf of a community[8]), academic informatics groups, or clinical information system vendors.

The leadership and governing authority for such a project do not necessarily reside within the same entity. For example, in the Real-time Outbreak and Disease Surveillance (RODS)[9] Winter Olympic deployment in Salt Lake City, the RODS Laboratory, located in Pittsburgh, acted as the project's Trusted Broker, an entity to which data providers agreed to send data for analysis and reporting.[10] The Trusted Broker handled data storage, analysis, and reporting under the auspices of a governing body comprised of (1) representatives of the data providers, (2) the State Epidemiologist of Utah, and (3) the director of the RODS laboratory. The surveillance project was led by medical informaticians and physicians from the Universities of Pittsburgh and Utah and representatives from both state and local health departments in Utah.

The experience with leadership and coalition to date can be summarized as a set of different possible models that vary by the scope of region, by who drives the project, and by whether that entity has legal authority to collect data.[11,12]

Special Event Model

In this model, teams of public health officials "drop in" to cover an event such as the 1999 World Trade Organization meeting in Seattle, the 2002 World Series in Phoenix, or the September 11th World Trade Center attacks.[13,14] Data are collected manually using special purpose forms from regional hospitals for the duration of the event. The legal authority is conferred by state or local public health statutes, which may be enacted specifically for the event. Regional health departments do much of the work with assistance often requested from the Centers for Disease Control and Prevention (CDC), independent contractors, and, in some cases, the military.[15] The drivers typically are local public health officials.

Regional Model

A region could be a state, a large city, a county or group of counties, or a small city. A population density that typically crosses local health jurisdiction boundaries defines the surveillance area. The technical work can be performed by any of a number of entities. Drivers may be a coalition of hospitals,[16] health care delivery organizations, county health departments,[12] or an informatics group.

Proposed Public Health Information Network (PHIN) Model

The geographic unit of organization is a state, comprised of a set of local health jurisdictions, each with primary responsibility for detection, investigation, and, at least in certain cases, management of disease outbreaks. The intended scope of coverage is the entire nation. The legal authority is state or local public health statutes. The state and local health departments develop systems internally or with the aid of contractors. CDC funding and guidance are the drivers for the PHIN project.[17,18]

Military Model

The scope of coverage in this model can be a region like the Washington, DC, National Capital area, which has a large military presence,[19] or the global military community with data coming from installations throughout the world.[20] Data are collected under the legal authority of the military. Although analogous to a civilian model, the military drives the project and does the work.

Selecting the Population and the Data

Data Sources

The geographic, demographic, and temporal coverage must be sufficient to support anomaly detection. The most valuable data sources will be those that are electronically stored, allow robust syndromic grouping, and are available in a timely fashion. Additional sources of data, such as electronic medical records that may not yet be in sufficiently widespread use today, may offer expanded opportunities in the near future. So far, practicality has dictated use of data already collected for other purposes. Implementing new data collection processes has a prohibitive cost, and the health care workers have repeatedly shown poor compliance with additional administrative tasks.[21] While data for other purposes may not be perfectly suited to the task of outbreak detection and monitoring, using them ensures availability of baseline data, which are valuable for algorithm development, and reduces the effort and costs associated with introducing new processes and software into existing workflows.

Identifying the Syndrome in the Population

Initially, the system developer must decide which diseases need to be detected and which syndromes, therefore, should be tracked. A data source can be chosen anywhere along the continuum of the disease process, and the types of data that have been used or considered are myriad. Citizenry may be observed, be polled, or have selected aspects of their public behavior analyzed. Behaviors of the citizenry, when their health is affected, may leave imprints on certain data sets. The principal underlying premise of these systems is that the first

signs of a covert biological warfare attack will be clusters of victims who change their behavior because they begin to become symptomatic (Fig. 1). When people become sick, they may make purchases such as facial tissues, orange juice, and over-the-counter remedies for colds, asthma, allergies, intestinal upsets, and so on. They may not report to school or work. Less traditional data sources include work and school absenteeism and retail sales [22] of groceries[23] and over-the-counter medication,[24] including electrolyte products for pediatric gastroenteritis.[25] The next level of detectable activity is likely to be encounters with the health care system. Patients may phone in to nurses or physicians. They may visit sites of primary care,[26] activate 911 emergency medical services,[27] visit emergency departments,[28,29] or be hospitalized. They may have laboratory tests ordered.[30] Some may die. All of this activity may precede the first confirmed diagnosis of a bioterrorism victim.

Acquiring and Organizing Data

Data Entry and Storage

Once the choices of population and data have been made, the next step is to acquire and manipulate the data. Data acquisition can be manual or automatic. Manual acquisition requires personnel resources of some kind—to cull a log, e-mail a report, or transfer a file, whenever data are to be transmitted. Automated processes may result in the transmission of a text report, a data file, or a series of structured messages over an error-tolerant interface but do not require human intervention to trigger each report. For all types of data, those that are already electronically coded in some format will be simpler to transfer and may provide information more rapidly.

If readily available data do not provide a clear picture of the health status of the community being monitored, new data can be collected from the surveillance system. Systems, including RSVP[31] and LEADERS,[15] have been developed using Web-based or handheld devices that allow providers to manually enter information at the time of patient care. These systems allow more specific and complete patient syndromic information to be gathered and would enable better identification of patients who have the condition of interest but face the challenge of provider acceptability and compliance. For example, when drop-in surveillance involving manual data entry was instituted in New York City around Ground Zero, data collection was difficult even with the infusion of short term, dedicated personnel. Afterward, the effort was unsustainable without outside assistance.[14]

Syndromic Grouping

Once the data have been identified and obtained, the next step is to logically group them in some way that provides useful information. While health care data sources often enable more fine-grained syndromic grouping (for example, respiratory illness, gastrointestinal illness), other data sources, such as school absenteeism, do not allow the assignment of each person into a syndromic category.

Developers of the first wave of syndromic surveillance systems have found that health care encounter data, and particularly emergency department data, are readily available and well suited to syndromic surveillance. Real-time

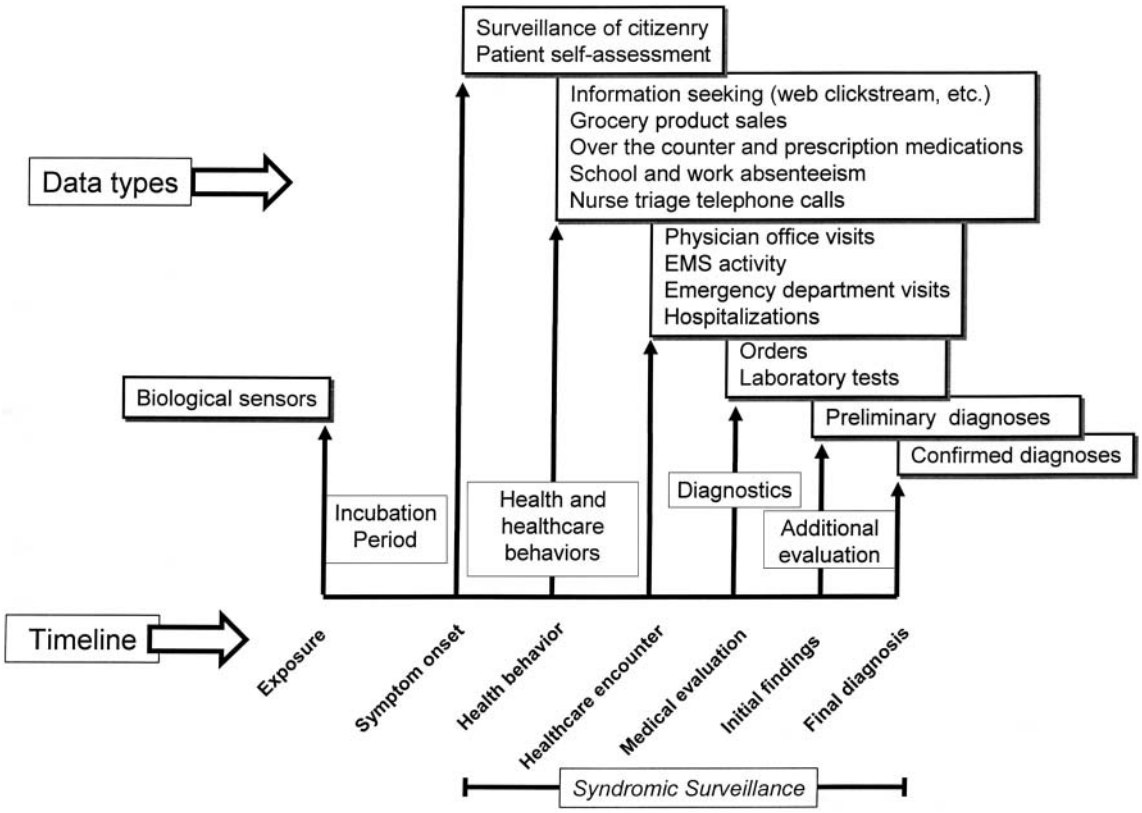

Figure 1. A progression of useful data sources as related to the underlying infection and associated behaviors.

data streams from these emergency department encounters have been established successfully in a number of regions.[11] Most emergency departments record patients' chief complaints at triage, and many do so electronically. Free-text chief complaints can be grouped into syndromes using tools such as the University of Pittsburgh CoCo Bayesian classifier, released as free software.[32,33] All U.S. emergency departments rely on the same standard for billing, the International Classification of Diseases, 9th Edition, Clinical Modification (ICD),[34] a disease classification designed for aggregating cases with similar diagnoses. Studies have found that chief complaints and/or ICD codes can be used to group emergency department encounters into syndromes.[32,35–38] Since at many institutions, ICD codes are often assigned to emergency department cases days or even weeks after an encounter, they are not consistently useful for real-time surveillance. However, evidence suggests that ICD codes may more accurately classify patients into syndromes than chief complaints,[36] and further, that using ICD codes in outbreak detection yields improved performance.[39]

Architects of the Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE) system[40] developed a mapping of ICD codes to syndrome categories,[20] which has been widely distributed (available at www.geis.ha.osd.mil). The original diagnostic groupings were determined a priori based on expert opinion. After ESSENCE had generated sufficient baseline data, actual ICD code usage was measured, allowing for modification of the code set. In developing the ESSENCE code groups, use of ICD codes during ambulatory encounters was evaluated. For example, using the yearly influenza season as a benchmark for the accuracy of syndromic ICD code groupings, it was found that codes for allergic conditions, e.g., allergic rhinitis, did not increase during influenza season but did during the spring and fall months, so these codes were excluded from the respiratory group. Conversely, while otitis media was not included in the original grouping, it did strongly correlate with the yearly outbreak and was added to the ESSENCE respiratory group.

Before using the standard code set at a new institution, however, it is important to be aware that there may be substantial interinstitutional variation in billing and coding practices. Therefore, it may be a good idea to evaluate standardized syndromic code groupings for each new data source and each new site. One method to accomplish this is to perform a chart review, with clinicians using standard criteria to assign each clinical encounter to a syndromic category.[32,35,36] Then, the sensitivity, specificity, and positive and negative predictive values of each code grouping can be measured using the chart review as a gold standard.

Integrating Data across Multiple Sites

If appropriate agreements have been made to share data, the next barriers to overcome are the technical ones.

Varying Syndromic Surveillance System Architectures

The hospital and clinical organizations that generate data use a multitude of different information systems, some designed internally, others from a wide variety of vendors. Further, there are diverse syndromic surveillance implementations and, correspondingly, a wide range of architectures. These include Health Level 7 (HL7) interfaces (both message driven and batch) and query-based systems (both using platform-dependent protocols such as ODBC and open protocols such as those based on XML). Most systems create and query a central data repository.

Data Standards

The disparate legacy systems in data-producing facilities typically do not use standard formats to store or transmit data. Integration and interpretation across multiple regions would be greatly facilitated by the universal adoption of standards. In the meantime, it will be necessary to develop translation engines that transform data from existing formats to standard formats. The initial wave of systems developers, recognizing this difficulty, limited the data they collected to types of data that did not have this type of problem. In fact, a survey of eight syndromic surveillance systems[11] showed a striking convergence of clinical data elements used, including age, gender, free-text chief complaint, ICD-9 coded discharge diagnosis, and some form of spatial location (most often zip code).

Important standards include the Logical Observation Identifier Names and Codes (LOINC),[41] an internationally accepted standard to identify results and observations. Whether referring to a laboratory value (potassium, white blood cell count), or a clinical finding (blood pressure, electrocardiogram [EKG] pattern), unique and unambiguous codes are available. The Unified Medical Language System (UMLS)[42] provides a cross reference among a number of different coding systems, and a semantic structure defining relationships among different clinical entities. The Systematized Nomenclature of Medicine (SNOMED)[43] not only provides granular diagnostic codes but also permits recording of component and related concepts. HL7[44,45] is the health care standard messaging format, used for transmitting information among information systems in a variety of clinical and administrative settings.

In addition to these existing health care industry standards, the public health community and CDC are creating standard definitions to characterize what findings and diagnoses will be of interest to the public health department. Laboratory test and result codes are mapped to nationally notifiable disease conditions. There are other standards relevant to clinical or syndromic data collection. The CDC and eHealth Initiative Public Private Collaboration[46] have developed implementation guides for public health reporting of chief complaint information using version 2.3.1 of HL7 Standard Protocol. The Frontlines of Medicine[47] Working Group has balloted standards for a chief complaint coding scheme and an XML-based triage data report and has proposed a standard for emergency department case reports.

The PHIN[18] specification includes not only format and content standards, but also guidance on software architecture, access management, and data dictionaries. The National Committee on Vital and Health Statistics is charged with selecting standards for use in Health Insurance Portability and Accountability Act (HIPAA) transactions. In addition, the Secretary of the Department of Health and Human Services has announced adoption throughout the Federal government of HL7, LOINC, and Digital Imaging and Communications in Medicine (DICOM).

Operational Challenges to Integration

Even within a single institution, grouping all pertinent clinical, laboratory, and administrative data into a specific health care encounter is a challenge. Patient tracking across a regional syndromic surveillance system is a particularly difficult task. There is no universal health identifier in the Unites States, making it difficult to identify a patient who moves between institutions. These patients may be double-counted. Further, many of the data sets will be completely de-identified or contain only aggregated frequency data, making the tracking of an individual patient impossible.

Privacy Protection

Because outbreak surveillance requires analysis of data from large numbers of individuals, sometimes including private information, the confidentiality of the data must be carefully protected. There is tension, however, between this requirement and the need to retain the ability to re-identify individuals to follow-up on cases that are identified. When reporting case-based data, even when the name and hospital number are removed, the inclusion of identifiers such as race, date of birth, and zip code allows the re-identification of substantial numbers of patients.[48,49] Discovering disease through geospatial cluster recognition may require detailed address information for geocoding.

HIPAA

The legal status of syndromic surveillance is governed in part by state law, while the obligations and reporting requirements of health care institutions are governed by the HIPAA privacy rule. HIPAA regulations allow health care delivery organizations to disclose data to public health officials but do not require it. The laws that govern public health data reporting vary widely state to state.

HIPAA defines different use cases for protected health care data. The relevant use cases include health care operations, research, and public health operations. Below, a number of likely use cases are described as they apply to different health care organizational structures. The implications for meeting HIPAA requirements are discussed in this context.

The first scenario involves a single hospital wishing to implement a system for internal disease surveillance. If this system were to use only routinely collected health care data and provide aggregate results to appropriate health care providers for normal operational use, such as forecasting staffing demand based on disease levels, this would constitute a "health care operations" use and no institutional review board (IRB) approval or other modifications for HIPAA would be necessary.

If the disease surveillance effort is a research project that uses patient-identifiable information, then IRB approval is required by the Federal Office for Human Research Protections. If none of the 18 individual personal identifiers enumerated in the HIPAA privacy rule[50] are stored, data could be released to researchers as a "limited data set," and, under these conditions, a data use agreement must be signed.

In the case of a single hospital system reporting surveillance data to public health authorities, the HIPAA privacy regulations permit the unencumbered transmission of such information if it meets the criteria for public health activity. An accounting of such disclosures may be required.[50] The HIPAA security regulations require methods of protecting the data in transport, such as data encryption, secure sockets, secure shell tunneling, or the use of a virtual private network.

Outbreak Detection

Whatever data are used, the goal of outbreak detection is to distinguish an abnormal pattern from a normal one. We explore methods for accomplishing this with temporal and spatial data.

Control Charts

Control chart approaches, such as the cumulative sum (CuSUM),[51] rely on cumulative differences between observed and expected data in a time window when compared with a threshold. In traditional CuSUM, the expected data are simply a theoretical mean, which is constant over time. A suspicious increase in the observed data over the theoretical mean is evidence for an emerging outbreak. To allow for sampling variability, the threshold of the maximum difference between observed and expected values is typically some multiple of the standard error of the sample mean. Because many health care data sets show regular periodicities—one example is in Figure 2, which shows the number of daily visits of patients with respiratory syndromes at the emergency department of Children's Hospital Boston between June 1992 and February 2003—the theoretical mean needs to change over time to reflect annual periodicities such as increasing hospital visit rates in winter. The CuSUM method was corrected for seasonal and daily variations and is implemented in the CDC's Early Aberration Reporting Systems (EARS).[52]

Temporal Modeling Approaches

Other approaches involve comparing observed patterns with those predicted by a model. This approach requires a robust model of the baseline pattern of syndromes as well as the selection of a threshold to signal an alarm. Threshold values

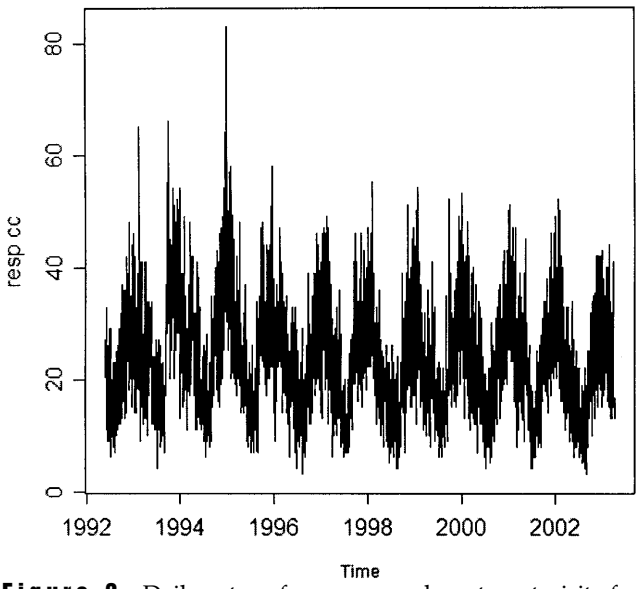

Figure 2. Daily rates of emergency department visits for respiratory syndromes as tracked by the AEGIS system at Children's Hospital Boston from 1992 until early 2003.

are a multiple of the standard error of the prediction. Typically, a value between 2 and 3.5 is chosen as the multiplier to ensure a false alarm rate below 5%.

To establish normal patterns, at least one or more years of historical data at the surveillance sites is required. These data will include regular recurrences of cyclic diseases such as influenza and local variations and trends in population density, hospital catchment areas, and shifting referral patterns. Typical models for temporal data are regression type models,[53] classical autoregressive integrated moving average (ARIMA) models,[54] or a combination of both methods. Serfling's method uses cyclic regression to model the normal pattern of the numbers of patients susceptible to death for pneumonia and influenza when there is not an epidemic with the objective of determining an epidemic threshold. Its use requires a clear definition of the disease, the selection of data to identify a normal pattern of susceptible patients, and the assumption that the normal pattern is periodical. Serfling's method has been adapted to model hospital visitation data for influenza.[55] In syndromic surveillance, the goal is to identify clusters of yet undiagnosed diseases, and the recurrent incidence of cyclic diseases should be part of the normal pattern of diseases that underlies, for example, the dynamics of hospital visit rates. Traditional ARIMA models seem better suited to describe historical visit rates and can account for temporal dependency, trends corresponding to secular changes in the populations, and seasonal effects.[29] Because a series of consecutive alarms can signify a real aberration rather than an unusual event, multiday temporal filters in which a weighted prediction of multiple days at once is compared with a threshold can lessen the effects of the large variability of hospital visit rates and improve both the timeliness and sensitivity of detection.[56] In the Automated Epidemiologic Geotemporal Integrated Surveillance (AEGIS) program at Children's Hospital Boston and Harvard Medical School, a hybrid of ARIMA with cyclic regression was found to have excellent predictive ability.

Another set of methods relies on Hidden Markov Models[57] to describe the normal pattern of diseases by using a hidden state that describes the presence or absence of an epidemic of a particular disease and a model of the data conditional on the epidemic status.[58] Closely related to Hidden Markov Models are change point algorithms to detect changes in a baseline model describing the normal pattern of hospital visits.[51,59] A common feature of the methods described is that they use aggregate data to model a normal pattern. However, these methods may be unable to detect small changes that affect only a specific group. The What's Strange About Recent Events system[60] is designed to complement traditional detection systems by looking for irregularities in the raw data. The system searches for irregularities in the data by using a set of rules and comparing the number of selected cases with the same number of cases recorded the week before.

Spatial and Spatiotemporal Modeling

Consideration of the spatial distribution of syndrome cases may facilitate the detection of a bioterrorism attack, particularly if the cases are distributed over space in a manner that is different from the background distribution. An initial consideration in conducting spatial surveillance is whether to use case point locations or counts of cases by regions. Use of case locations is generally preferable, as aggregation of cases to region counts tends to result in a loss of precision. At Children's Hospital Boston and the Harvard School of Public Health, new methods for geospatial cluster detection rely on the recognition of perturbations in the distribution of pairwise distances among all individual cases in a geographical area; this approach yields substantial power for detection and is used in the AEGIS program.[61]

However, there are many hurdles to overcome to use geographic location in surveillance. For instance, the only address that tends to be available in hospital information systems is the home address, and exposures may occur elsewhere. Second, there are privacy concerns when using the exact street address for each surveillance record. Third, considerable error occurs in the process of geocoding street addresses.[62] Finally, interpolating covariates to the case locations is difficult. If one is using region counts for surveillance, the first problem to be addressed is the selection of the regions. It is well known that scale (the number of regions for a given area) and zoning (the partitioning of a given area into the number of regions) can both affect the degree to which spatial processes can be detected.[63]

Spatial analysis can be incorporated into surveillance in a number of ways. The most simple approach is to examine the spatial distribution of observed cases or case counts over a fixed time interval without respect to time. A variety of methods are available to assess case location and region count clustering in general,[64] clustering at specific locations, [65] and clustering in relation to putative point sources.[66] While there is not explicit consideration of time in these approaches, they are implemented easily, and it is possible to informally compare results across different time intervals. A more powerful approach is to examine the joint spatial and temporal distribution of case locations or case counts over a fixed time interval. The method devised by Knox[67] and extended by Mantel[68] enables detection of space–time interaction in case locations compared with control locations. However, both of these methods require a priori selection of spatial and temporal distance parameters. Space–time scan statistics[69,70] avoid these assumptions and are useful to identify "suspect clusters" of case locations or region counts by using a window that moves in time and space. A desirable approach is to sequentially examine the joint spatial and temporal distribution of case locations or case counts over a dynamic temporal interval.

Many of these methods are still under development or being adapted to the context of syndromic surveillance. Some software to accomplish some of these tasks is available publicly , including the RODS outbreak detection software[33] and the SaTScan software.[71]

Measuring Surveillance System Quality

Of the important characteristics of public health surveillance systems,[72] three are especially important for the evaluation of syndromic surveillance systems: sensitivity, specificity, and timeliness. Developers should use these metrics to understand data quality and timeliness as well as more difficult questions such as which outbreaks can be detected, how large they must be to be detected, and how early they can be detected.

Data Quality

The term *data quality* refers to the accuracy of data and is a generic term not limited to public health surveillance.[73] The standard method for characterizing data quality measures the sensitivity and specificity with which the data can accurately classify patients relative to a criterion determination (gold standard). A critical design decision in such studies involves the criterion classification. If the criterion classification is too broad (e.g., includes cases of chronic respiratory illness), a misleadingly high sensitivity can be reported.

Timeliness

Timeliness refers to the time when a datum of interest becomes available relative to the time of occurrence of some reference event, such as the time of presentation of a patient to an emergency department. It is not always possible to measure timeliness. For example, if the data are not personally identifiable (i.e., over-the-counter sales of grocery products), they cannot be linked to a reference event. In such cases, timeliness may be calculated through aggregate measures, for example, sales of over-the-counter cough products begin to rise relative to when rates of emergency room visits for influenza begin to rise. Timeliness also has been estimated by studies of the behavior of sick individuals.[74]

Impact on Outbreak Detection

It is important to note that data quality does not have to be perfect for successful detection of disease outbreaks. In fact, the very earliest detection will likely come from statistical analysis of noisy data—for example, over-the-counter sales of medications—rather than from highly accurate, but late data such as microbiology culture results. Therefore, a potential data source should be judged by the combination of its data quality and timeliness as well as knowledge of the cost of false alarms versus the cost of delays in triggering true alarms for a specific disease threat.[75]

The term *outbreak detection performance* refers to the direct measurement of sensitivity, specificity, and timeliness of detection of outbreaks.[76] Such studies are, however, difficult to conduct due to the low frequency or even absence (e.g., smallpox) of outbreaks of most diseases. There is such difficulty in conducting direct analyses of outbreak detection performance, that relatively few studies are available, and those that exist typically have small sample sizes[77] (e.g., one outbreak) or simulations.[55] We recommend, however, that developers pay particular attention to the results of such studies as they become available because they will represent the most direct and rigorous determinations of the ability to detect outbreaks in real time using syndromic data.

Integration of Syndromic Surveillance with Public Health Response

If syndromic surveillance is to fulfill its goal of early outbreak detection, it must be linked tightly and integrally to medical care and to public health investigation and response. Syndromic surveillance relies on nondiagnostic data and monitoring of nonspecific signs and symptoms. These syndromic "signals" are akin to the alarming of a smoke detector. In most cases, the smoke is caused by burning toast, but each alarm must be investigated if fires are to be averted. In New York City, results of syndromic analyses are examined every day by analysts and a medical epidemiologist, and field teams are available for investigation and response 365 days a year, although they are rarely used.

Public Health Investigations

In conducting a public health investigation, the first task is to differentiate natural (statistical) variability as well as "pseudo-outbreaks" due to data entry or coding errors from a true increase in (infectious) illness. To some extent, "drilling down" into the available data can do this, especially if there are individual-level data available (as opposed to counts), or if clinical information systems can be queried in real-time. Lack of corroboration from other syndromic data sources can also be comforting. Finally, if the observed increase is not sustained in the next period of observation, then it is an important clue that this may be an artifact or normal statistical variability.

If an increase in syndromic events is thought to reflect a true increase in illness, then the next task is to differentiate self-limited natural illness from infectious disease outbreaks of public health significance, including bioterrorism. Suspicion may be increased if the profile of the cases is unusual in their geographic distribution (spatially clustered), demographics, or symptoms. But this investigation will require telephone calls at a minimum and possibly on-site investigations, including active surveillance for more severe manifestations. Clinicians and medical examiners may need to be interviewed. Another approach has been to follow-up on individuals who formed the cluster resulting in a syndromic surveillance signal. These patients or their physicians can be contacted and asked about any deterioration in their medical condition, unusual manifestations of illness, or shared exposures. But, ultimately, if early diagnosis of a bioterrorist attack is realized, it will be made through obtaining diagnostic laboratory or radiologic studies on individuals with mild illness who otherwise would have probably not received these studies. Communication with front-line medical personnel and heightening their clinical "prior probability" for recognizing the prodrome of a severe illness is a necessary part of this phase of the response. A patient with flulike symptoms who is presenting to an emergency department that is located in an area with a suspicious respiratory signal might be treated with more caution, not unlike the attention given to postal workers with "flu" after October 2001.

Synergies

Syndromic surveillance programs that are integrally linked to public health response also benefit tangibly from this relationship. The competing priorities of public health will ensure that systems have multiple uses (monitoring regional patterns of asthma and gastrointestinal outbreaks as well as bioterrorism), and do not have unrealistically high rates of false alarms. To fulfill the overarching mandate of early detection, systems will be built to utilize data that are available "real-time" 365 days per year, rather than data that function admirably on retrospective data analysis but are not available on weekends or holidays or are associated with a 48-hour lag.

Second, being linked to public health response allows system developers to learn from prospective experience. If routine signals are not investigated, there is no opportunity to validate the data sources and algorithms in the real world

or to improve the ability of systems to differentiate true infectious disease clusters from false alarms.

Alarm Thresholds

Finally, alarm thresholds should be set based on explicit utility considerations that attempt to optimize the tradeoff between the cost of false alarms and the expected benefits of earlier detection. In the aftermath of the anthrax mail attacks, the Bayesian "prior probability" of a massive aerosolized anthrax attack on New York City in the next 30 days was dramatically heightened, and public health resources were mobilized and on high alert. Operators of a detection system in this situation might set the detection threshold lower to achieve earlier detection at the cost of frequent investigations of false alarms.

Next Steps

Syndromic surveillance system developers face several challenges that can be addressed through rigorous research. Designing "dual use" systems will boost sustainability. If a surveillance system is designed to only detect bioterrorism or very rare outbreaks, its use and funding allocation will diminish over time if there are no events. However, if the system is designed to help clinicians, public health officials, and researchers automate existing data collection processes and provide new streams of data, then it is more likely to be maintained, improved, and used. Furthermore, it is more likely to be up and running should a bioterrorist attack occur.

Optimal data sources for surveillance must be identified and thoroughly assessed. Syndrome definitions that lead to high performance outbreak detection must be developed and assessed. Privacy-preserving data integration methods must be developed, formalized, and implemented.

Syndromic surveillance systems can now be trained on data sets that include naturally occurring outbreaks. Since data on bioterrorism attacks ARE extremely limited, none of the detection algorithms can be trained on real data sets for the purpose of bioterrorism detection. Therefore, realistic simulation is necessary, possibly requiring development of detailed attack scenarios. To benchmark the performance of detection and monitoring systems, training and validation data containing signal and noise are required. These data can be samples of authentic regional data, synthetic data, or a combination of both (semisynthetic data). The global ability of a system to detect "bioterrorism" cannot be assessed. Rather, performance at detecting attacks with specific agents under specific conditions needs to be measured. Metrics for system performance have been proposed in the CDC draft guidelines for evaluation of syndromic surveillance systems.[76] A rigorous method for evaluation is the receiver operating characteristic (ROC) curve. This method involves plotting sensitivity against (1 minus the specificity) and it allows comparisons without any assumptions about detection thresholds, effectively comparing outbreak detection performance at all operational settings simultaneously. In addition, there is a need for detection methods that formally integrate multiple disparate data sources over space and time.[78]

Conclusion

Traditional surveillance and astute clinicians will always play a critical role in the accurate diagnosis and treatment of patients as well as in the identification of public health emergencies. However, syndromic surveillance is another modality that clearly has a role in detecting and monitoring bioterrorism as well as other outbreaks and public health problems. The work to be done over the coming months and years is to build our data integration infrastructure, develop and refine our methods, and estimate, to the best of our ability, the promise and limits of our technology.

References ■

1. Centers for Disease Control Prevention. Biological and chemical terrorism: strategic plan for preparedness and response. Recommendations of the CDC Strategic Planning Workgroup. MMWR. 2000;49(RR-4):1–26.
2. Jernigan JA, Stephens DS, Ashford DA, et al. Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States. Emerg Infect Dis. 2001;7:933–44.
3. Kohane IS. The contributions of biomedical informatics to the fight against bioterrorism. J Am Med Inform Assoc. 2002;9:116–9.
4. Teich JM, Wagner MM, Mackenzie CF, Schafer KO. The informatics response in disaster, terrorism, and war [comment]. 2002:97–104.
5. Koplan J. CDC's strategic plan for bioterrorism preparedness and response. Public Health Rep. 2001;116(suppl 2):9–16.
6. Yasnoff WA, Overhage JM, Humphreys BL, et al. A national agenda for public health informatics. J Public Health Manage Pract. 2001;7(6):1–21.
7. O'Toole T. The problem of biological weapons: next steps for the nation. Public Health Rep. 2001;116(suppl 2):108–11.
8. Gesteland PH, Wagner MM, Chapman WW, et al. Rapid deployment of an electronic disease surveillance system in the state of Utah for the 2002 Olympic winter games. Proc AMIA Symp. 2002:285–9.
9. Tsui F-C, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical description of RODS: a real-time public health surveillance system. J Am Med Inform Assoc. 2003;10:399–408.
10. Gesteland PH, Gardner RM, Tsui F-C, et al. Automated syndromic surveillance for the 2002 winter Olympics. J Am Med Inform Assoc. 2003;10:547–54.
11. Lober WB, Karras BT, Wagner MM, et al. Roundtable on bioterrorism detection: information system-based surveillance. J Am Med Inform Assoc. 2002;9:105–15.
12. Mostashari F, Hartman J. National Syndromic Surveillance Conference. New York Academy of Medicine [Web posting of conference proceedings]. Available at: http://www.nyam.org/events/syndromicconference/. Accessed Jan 8, 2003.
13. Das D, Weiss D, Mostashari F, et al. Enhanced drop-in syndromic surveillance in New York City following September 11, 2001. J Urban Health. 2003;80(suppl 1):I76–I88.
14. Centers for Disease Control and Prevention. Syndromic surveillance for bioterrorism following the attacks on the World Trade Center–New York City, 2001. MMWR. 2002;51(Spec No):13–5.
15. Green MS, Kaufman Z. Surveillance systems for early detection and mapping of the spread of morbidity caused by bioterrorism. Harefuah. 2002;141(Spec No):31–3, 122.
16. McDonald C, Overhage J, Dexter P, et al. The Regenstrief medical record system. Proc AMIA Symp. 1999:1212.
17. National Electronic Disease Surveillance System Working G. National Electronic Disease Surveillance System (NEDSS): a standards-based approach to connect public health and clinical medicine. J Public Health Manage Pract. 2001;7(6):43–50.
18. Centers for Disease Control and Prevention. Public Health Information Network goals. Available at: http://www.cdc.gov/phin/about/goals.htm. Accessed Jul 11, 2003.
19. Lewis MD, Pavlin JA, Mansfield JL, et al. Disease outbreak detection system using syndromic data in the greater Washington DC area. Am J Prev Med. 2002;23:180–6.

20. DoD-GEIS. Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE). Available at: http://www.geis.ha.osd.mil/GEIS/SurveillanceActivities/ESSENCE/ESSENCEinstructions.asp. Accessed Oct 28, 2003.
21. Committee on Quality of Health Care in America Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press, 2001.
22. Wagner MM, Robinson JM, Tsui F-C, Espino JU, Hogan WR. Design of a national retail data monitor for public health surveillance. J Am Med Inform Assoc. 2003;10:409–18.
23. Reference deleted.
24. Goldenberg A, Shmueli G, Caruana RA, Fienberg SE. Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. Proc Natl Acad Sci U S A. 2002; 99:5237–40.
25. Hogan WR, Tsui F-C, Ivanov O, et al. Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products. J Am Med Inform Assoc. 2003; 10:555–62.
26. Lazarus R, Kleinman K, Dashevsky I, et al. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. Emerg Infect Dis. 2002;8:753–60.
27. Greenko J, Mostashari F, Fine A, Layton M. Clinical evaluation of the emergency medical services (EMS) ambulance dispatch-based syndromic surveillance system, New York City. J Urban Health. 2003;80(suppl 1):I50–6.
28. Lober WB, Trigg LJ, Karras BT, et al. Syndromic surveillance using automated collection of computerized discharge diagnoses. J Urban Health. 2003;80(suppl 1):I97–I106.
29. Reis BY, Mandl KD. Time series modeling for syndromic surveillance. BMC Med Inform Decis Making. 2003;3(1):Available at: http://www.biomedcentral.com/1472-6947/1473/1472.
30. Jernigan DB, Koski E, Shoemaker HA, Mallon RP, Pinner RW. Evaluation of laboratory test orders and results in a national laboratory data repository: implications for infectious diseases surveillance [abstract]. International Conference on Emerging Infectious Diseases, Jul 2000, Atlanta, GA.
31. Zelicoff A, Brillman J, Forslund DW, et al. The Rapid Syndrome Validation Project (RSVP). Proc AMIA Symp. 2001:771–5.
32. Ivanov O, Wagner MM, Chapman WW, Olszewski RT. Accuracy of three classifiers of acute gastrointestinal syndrome for syndromic surveillance. Proc AMIA Symp. 2002:345–9.
33. RODS Laboratory. RODS: Real-time Outbreak and Disease Surveillance Software Packages. Available at: http://www.health.pitt.edu/rods/sw/default.htm. Accessed Jan 31, 2003.
34. American Medical Association. ICD-9-CM 2002: International Classification of Diseases, Volumes 1 and 2. 9th Revision ed; 2002.
35. Espino JU, Wagner MM. Accuracy of ICD-9-coded chief complaints and diagnoses for the detection of acute respiratory illness. Proc AMIA Symp. 2001:164–8.
36. Beitel AJ, Olson KL, Reis BY, Mandl KD. Use of emergency department chief complaint and diagnostic codes for identifying respiratory illness in a pediatric population. Pediatr Emerg Care. 2004. (in press).
37. Aronsky D, Kendall D, Merkley K, James BC, Haug PJ. A comprehensive set of coded chief complaints for the emergency department. Acad Emerg Med. 2001;8:980–9.
38. Begier E, Sockwell D, Branch L, et al. The National Capitol Region's emergency department syndromic surveillance system: do chief complaint and discharge diagnosis yield different results? Emerg Infect Dis. 2003;9:393–6.
39. Reis BY, Mandl KD. Syndromic surveillance: the effects of syndrome grouping on outbreak detection performance. Ann Emerg Med. 2004. (in press).
40. Lombardo J, Burkom H, Elbert E, et al. A systems overview of the electronic surveillance system for the early notification of community-based epidemics (ESSENCE II). J Urban Health. 2003;80(suppl 1):I32–42.
41. Logical Observation Identifiers Names and Codes (LOINC®). Available at: http://www.loinc.org. Accessed Jan 30, 2003.
42. National Library of Medicine. Unified Medical Language System (UMLS). Available at: http://www.nlm.nih.gov/research/umls/. Accessed Jan 31, 2003.
43. SNOMED. Available at: www.snomed.org. Accessed Jan 30, 2003.
44. Beeler GW. HL7 version 3–an object-oriented methodology for collaborative standards development. Int J Med Inf. 1998; 48(1–3):151–61.
45. Health Level Seven. Available at: http://www.HL7.org. Accessed Jan 30, 2003.
46. eHealth Initiative. Available at: http://www.ehealthinitiative.org. Accessed Jan 30, 2003.
47. Barthell EN, Cordell WH, Moorhead JC, et al. The Frontlines of Medicine Project: a proposal for the standardized communication of emergency department data for public health uses including syndromic surveillance for biological and chemical terrorism. Ann Emerg Med. 2002;39:422–9.
48. Sweeney L. Replacing personally-identifying information in medical records, the Scrub system. Proc AMIA Annu Fall Symp. 1996:333–7.
49. Sweeney L. Guaranteeing anonymity when sharing medical data, the Datafly System. Proc AMIA Annu Fall Symp. 1997: 51–5.
50. Centers for Disease Control and Prevention. HIPAA privacy rule and public health. Guidance from CDC and the U.S. Department of Health and Human Services. MMWR. 2003;52(suppl):1–17, 19–20.
51. Basseville M, Nikiforov IV. Detection of Abrupt Changes: Theory and Application. Saddle River, NJ: Prentice Hall, 1993.
52. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). J Urban Health. 2003;80(2 suppl 1):i89–96.
53. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths. Public Health Rep. 1963;78: 494–506.
54. Box GEP, Jenkins GM. Time Series Analysis: Forecasting and Control (ed 2). San Francisco, CA: Holden-Day, 1976.
55. Tsui FC, Espino JU, Wagner MM, et al. Data, network, and application: technical description of the Utah RODS Winter Olympic Biosurveillance System. Proc AMIA Symp. 2002:815–9.
56. Reis BY, Pagano M, Mandl KD. Using temporal context to improve biosurveillance. Proc Natl Acad Sci U S A. 2003;100: 1961–5.
57. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE. 1989;77:257–85.
58. Le Strat Y, Carrat F. Monitoring epidemiologic surveillance data using hidden Markov models. Stat Med. 1999;18:3463–78.
59. Sebastiani P, Ramoni M, Cohen P. Bayesian clustering by dynamics. In: Sun LgaR (ed). Sequence Learning: Paradigms, Algorithms, and Applications. New York, NY: Springer; 2001, pp 11–34.
60. Wong W, Moore A, Cooper G, Wagner M. Rule-based anomaly pattern detection for detecting disease outbreaks. Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02), 2002, pp 217–23.
61. Olson KL, Bonetti M, Pagano M, Mandl KD. Enhanced power to detect bioterrorism with spatial clustering [abstract]. Pediatr Res. 2002;51(4 pt 2):132a.
62. Krieger N, Waterman P, Lemieux K, Zierler S, Hogan J. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. Am J Public Health. 2001;91:1114–6.

63. Fotheringham A, Wong D. The modifiable areal unit problem in multivariate statistical analysis. Environment and Planning A. 1991;23:1025–44.
64. Cliff A, Ord J. Spatial Processes: Models and Applications. London: Pion Ltd., 1981.
65. Anselin L. Local indicators of spatial autocorrelation—LISA. Geograph Anal. 1995;27:93–115.
66. Waller LA. A civil action and statistical assessments of the spatial pattern of disease: do we have a cluster? Regul Toxicol Pharmacol. 2000;32:174–83.
67. Knox EG. Detection of space-time interactions. Appl Stat. 1964;13:25–30.
68. Mantel N. The detection of disease clustering and a generalised regression approach. Cancer Res. 1967;27:209–20.
69. Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. J R Stat Soc. 2001;164 (part 1):61–72.
70. Kulldorff M, Athas W, Feurer E, Miller B, Key C. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am J Public Health. 1998;88: 1377–80.
71. National Cancer Institute. SaTScan™–Spatial and Space-Time Scan Statistics. Available at: http://srab.cancer.gov/satscan/download.html. Accessed Jan 31, 2003.
72. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems. MMWR. 2001;50(RR13):1035.
73. Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. J Am Med Inform Assoc. 1997;4:342–55.
74. Zeng X, Wagner MM. Modeling the effects of epidemics on routinely collected data. Proc AMIA Symp. 2001:781–5.
75. Wagner MM, Tsui FC, Espino JU, et al. The emerging science of very early detection of disease outbreaks. J Public Health Manage Pract. 2001;7(6):51–9.
76. Sosin DM. Draft framework for evaluating syndromic surveillance systems. J Urban Health. 2003;80(suppl 1):I8–13.
77. Tsui FC, Wagner MM, Dato V, Chang CC. Value of ICD-9 coded chief complaints for detection of epidemics. Proc AMIA Symp. 2001:711–5.
78. Reis BY, Mandl KD. Integrating syndromic surveillance data across multiple locations: effects on outbreak detection performance. Proc AMIA Symp. 2003:549–53.